

THE
UNIVERSITY
OF RHODE ISLAND

University of Rhode Island
DigitalCommons@URI

Environmental and Natural Resource Economics
Faculty Publications

Environmental and Natural Resource Economics

2014

Searching for the determinants of climate change interest

Patrick Cavanagh
University of Rhode Island

Corey Lang
University of Rhode Island, clang@uri.edu

See next page for additional authors

Creative Commons License



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).

Follow this and additional works at: https://digitalcommons.uri.edu/enre_facpubs

Citation/Publisher Attribution

Patrick Cavanagh, Corey Lang, Xinran Li, Haoran Misao, and John David Ryder. "Searching for the Determinants of Climate Change Interest." *Geography Journal*, vol. 2014, Article ID 503295, 11 pages.
Available in: <http://dx.doi.org/10.1155/2014/503295>

This Article is brought to you for free and open access by the Environmental and Natural Resource Economics at DigitalCommons@URI. It has been accepted for inclusion in Environmental and Natural Resource Economics Faculty Publications by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons@etal.uri.edu.

Authors

Patrick Cavanagh, Corey Lang, Xinran Li, Haoran Miao, and John David Ryder

Searching for the determinants of climate change interest

Patrick Cavanagh, Corey Lang^{*}, Xinran Li, Haoran Miao, and John David Ryder
Department of Environmental and Natural Resource Economics
University of Rhode Island

September 10, 2014

Abstract

A meaningful CO₂ mitigation policy is unlikely at the national level in the United States. What is currently happening and what is much more likely to occur in the future is city and regional level efforts of mitigation and adaptation. This paper aims to understand the geographic and socioeconomic characteristics of metropolitan areas and regions that lead to engagement with the issue of climate change. We use geographically explicit, internet search data from Google to measure information seeking behavior, which we take to translate into engagement, attention and interest. Our spatial hotspot analysis creates a map that potentially could be harnessed by policymakers to gauge mitigation support or adaptation potential. The results of our multivariate analysis suggest that socioeconomic factors are the strongest determinants of search behavior and that climate and geography have little to no impact. With regard to political ideology, we find evidence of a non-linear, inverse-U relationship with maximum search activity occurring in metropolitan areas with a near even political split, suggesting parity may be good for engagement.

Keywords: climate change, adaptation, political ideology, attention, Google Trends

^{*}Corresponding author. 214 Coastal Institute, 1 Greenhouse Rd., Kingston, RI, 02881. Email: clang@mail.uri.edu. The authors declare that there is no conflict of interests regarding the publication of this paper. This paper is a contribution of the Rhode Island Agricultural Experiment Station (#5410).

1. Introduction

The United States has failed to achieve meaningful carbon legislation to limit emissions and mitigate damages. For many Americans and certainly politicians, the whole idea of climate change remains contentious, and therefore a path leading to meaningful legislation is difficult to envision. Despite the stasis at the federal level, cities and regions are finding the political will to take action on both mitigation and adaptation. Nine states in the Northeast have banded together in the Regional Greenhouse Gas Initiative (RGGI) and implemented a cap-and-trade policy to reduce emissions from electricity generation. Similarly, California recently enacted the Global Warming Solutions Act (AB 32) that will use market mechanisms to reduce all emissions to 1990 levels by 2020. New York City's PlaNYC and Boston's Climate Action Plan focus both on reducing emissions and improving resiliency in the face of rising sea levels and other climate change impacts. These local and regional actions are consistent with the ideas of Kahn [1], who argues that cities will drive adaptation in an effort to compete for footloose residents. Thus, understanding the characteristics of metropolitan areas that make them more or less likely to engage in climate change policy is critically important.

This paper seeks to add to that understanding by utilizing internet search data from Google Trends. Specifically, we examine how the characteristics of metropolitan areas, in terms of demographics (income, education, race), political preferences, geography, climate, and carbon intensity of their economy, determine search volumes for climate change-related words. We downloaded data from Google Trends on search activity for every media market (roughly equivalent to a metropolitan area) in the Continental United States (a total of 205). Data were gathered for the search terms "global warming" and "climate change", as well as "weather", which has different motivations behind it, but still serves as a baseline comparison and validity check for the data and methods.

Traditionally, researchers have gathered data on beliefs and interest level through survey techniques, and there are two reasons why internet search data offer benefits over survey data. First, a survey is a stated preference, where the respondent may or may not tell the truth. In contrast, internet searches represent a type of revealed preferences. The internet is private and anonymous, so there is no reason for searches to conform to existing notions of what type of person (e.g., conservative) should believe in climate change. Second, when a person uses Google to seek information, this is active engagement with a subject, which is far more important for

learning about and adapting to climate change than answering a yes or no question when someone calls you on the phone. The major drawback of internet search data is that they do not represent a clear opinion. We discuss in more detail how we view inference using Google Trends and how others have used this tool in Section 2.

Our first analysis examines search activity in a spatial context and seeks to determine if there are clusters of high or low search interest. The intent is to begin a process of providing policymakers and environmental activists with a map indicating regions where mitigation support may be high or low or prudent adaptation likely or unlikely to occur. The results indicate that clusters of high interest, or hotspots, are the Northeast, the Great Lakes region of the Midwest, California, and the Pacific Northwest. The Northeast and California already have enacted mitigation policies (RGGI, AB32), which generates confidence in the method and data and further may suggest that the Great Lakes region and Pacific Northwest are amenable to similar policies. The South, stretching from Texas to Alabama, is a cluster of low interest. This finding is perhaps surprising given the recency of Hurricane Katrina and projected increase in hurricane intensity and may suggest that education and outreach efforts are needed in this area.

Second, we analyzed the determinants of search behavior using multivariate regression with socioeconomic, geographic, and climate characteristics as independent variables. Our results suggest that socioeconomic factors are much stronger determinants of climate change-related searches than climate, geography, and CO₂ emissions. Specifically, we find that climate change-related searches are positively correlated with a metropolitan area's proportion of college graduates, proportion of Democrats, and proportion of whites and negatively correlated with a metropolitan area's average income. Metropolitan areas on the coast are no more or less likely to seek information about climate change, nor are hot or cold climate metropolitan area, nor areas with high or low CO₂ emissions per capita. Our findings contrast to those of O'Connor et al. [2], who argue that demographic factors tend to have modest relationships with climate change perceptions. Prior research has found a strong link between personal experiences of local weather and attitudes towards the legitimacy of climate change [3-4]. While we do not examine the effect of weather, our results add to this literature by showing that climate does not appear to influence search behavior.

Given the prominence of politics in debates over climate change, we examine in some detail how political affiliation affects internet search behavior related to climate change. While many

surveys have shown a substantial gap between democrats and republicans in the belief in climate change and willingness to take action [5], we speculated that this result could be due to posturing to project what political elites would expect respondents to say and using private search data could reveal a different pattern. We perform a semi-parametric analysis and find that for the search term “global warming”, search volume is greatest when metropolitan areas are more centrist, the peak occurs around 43% Democrat, with lower search volumes as metropolitan areas become more partisan. For the search term “climate change”, however, search activity increases as metropolitan areas lean more Democratic over the whole range of the political spectrum. These results first suggest that “climate change” may be more of a branded political word than “global warming”. Second, with “global warming”, our results may indicate that metropolitan areas that are more moderate are seeking information and engaging in healthy debate.

The main contribution of this paper is to add to the strong literature that seeks to understand beliefs of Americans about climate change [6-9]. By using a novel data set, we are able to have far more geographic detail than typical surveys, which allows us to expand on the spatial variation in climate change perception found by Hamilton and Keim [10] and examine correlation with a broader range of geographic and socioeconomic factors. Further, the data represent revealed preferences and active engagement, which are far more valuable when assessing learning and adaptive capacity.

2. Benefits and Considerations of Internet Search Data

For many Americans, the internet has become the primary means of gathering information. Among American internet users across all age groups, 90% utilize search engines for information seeking, and 83% of these queries are entered using Google Search [11]. Search engines give people considerable freedom to search for myriad topics without many of the limits and barriers associated with traditional information seeking methods.

Google Trends data are a measure of public attentiveness, engagement, interest, and issue salience. Internet search data have been used to: gauge religious influence within specific geographic areas [12]; determine if ballot measures were more likely to be voted on by voters [13]; measure Hurricane information seeking behavior following tropical storm events [14]; predict unemployment and sales [15]; and forecast cinema admissions [16]. Internet searches

have also been popular in the field of epidemiology, particularly in tracking the spread of the flu virus [17-18].

Beyond measuring interest, internet search data are a revealed preference alternative to traditional survey data and can be used to gather information on taboo or controversial topics. Stephens-Davidowitz [19] uses Google Trends to measure racist attitudes and improve 2008 presidential voting predictions. Most related to the present research, Kahn and Kotchen [20] and Lang [21] use Google data to investigate interest in the politically charged topic of global warming. Kahn and Kotchen [20] find that online interest in global warming declines as unemployment rises, and interestingly the decline is larger for Democrats. Lang [21] finds that weather fluctuations cause people to seek information about climate change and that different types of weather events lead to different responses in areas of different political ideologies and average educational attainment.

Additionally, Google Trends provide reliable, near real-time updates [22-23], which is prohibitively costly for the vast majority of surveys. Further, internet search data can more effectively gather indicators from underrepresented groups [24] and avoid interviewer bias [25].

Most of the research to date using Google Trends has focused on time variation and prediction. While there are advantages to a time series or panel data approach, often a substantial portion of variation in search is not used. For example, for the search term “global warming” over our sample period, 65% of variation is explained by media market and month-year fixed effects (dummy variables) alone, leaving only 35% of variation to identify a relationship. The purpose of this paper is to investigate how geography and spatial relationships may explain some of that variation.

Despite the many advantages over surveys, internet search data are fundamentally different than opinions and do not necessarily correlate with support for an issue.¹ For example, not everyone searching “climate change” would support a mitigation policy, but the data can still provide useful trends [27]. Search data is an indicator of an area’s willingness to invest some time and effort into actively executing a search engine query and consciously processing the

¹ Another concern with Google Trends is the extent to which it can represent the general population to the same degree as survey takers. The research examining this concern has been done in developing countries, where survey data are more prevalent and internet access is limited to a much smaller proportion of the population than in the United States, seemingly increasing the burden on search data. Carrière-Swallow and Labbé [26] and Mellon [22] validate search data as an equal or better indicator of a broad range of topics.

results [23]. Using search data to gauge this motivated search behavior has proven to be a valuable tool for researchers and will likely become increasingly popular as additional applications are explored and validated.

Turning the attention specifically to our focus of examining search data for “climate change”, “global warming”, and “weather”, different search terms can be classified by the user’s intention [28-29]. “Climate change” and “global warming” are classified as *informational*, people search these to learn more about a concept. In contrast, “weather” is classified as *resource* or *navigational*, that is people search this to obtain a quick fact (e.g., the daily forecast) or navigate to a website like weather.com, as opposed to something larger like wanting to understand the science behind meteorology. Understanding this classification is critical for the present study because it implies that the data from Google Trends are an aggregation of individuals seeking information and wanting to engage with the issue.

3. Data and Spatial Analysis

This section first describes the Google Trends data we collected. Second, we present maps that illustrate the degree of spatial variation in search activity and a Hot Spot analysis that looks for spatial clustering of search activity. Finally, we describe the extensive geographic and socioeconomic data collected for use in our multivariate regression analysis of the determinants of search activity.

3.1 Google Trends

We downloaded internet search activity in the United States using Google Trends for the keywords ‘global warming’, ‘climate change’ and ‘weather’. The smallest geographic unit that Google Trends offers for the U.S. is media market. Media markets consist of several contiguous counties, are typically associated with a metropolitan area, and are used primarily to gauge television viewership habits. The U.S. is broken into 210 media markets, however this paper focuses on the 205 media markets in the Continental U.S. We will refer to media markets as metropolitan areas, interchangeably.

Google Trends reports a time series index of search activity on a 0-100 scale, with 100 being the maximum search interest observed over that time and all other numbers being proportional to that maximum. Google also reports an average of this index, which can be interpreted as the

average search interest over time, and we refer to it as the average query share (AQS). However, at the moment, AQSs for each media market cannot be compared because they are all relative to their own maximum. Because we are interested in across space comparisons, we downloaded each media market's AQS simultaneously with the AQS for Washington DC. Now each media market's AQS is relative to Washington DC. We create our key variable Google Search Index (GSI), which will become the dependent variable in our multivariate regression, using the following formula:

$$GSI_i = AQS_{i,DC} \times \left(\frac{AQS_{DC,DC}}{AQS_{DC,i}} \right) \quad (1)$$

where $AQS_{i,j}$ is average query share for media market i with respect to media market j . The term in parentheses is a necessary scalar in case Washington DC does not have the maximum search interest compared to i . If Washington DC does have the maximum search interest, then the second term equals one. We restricted the time window of Google searches to be January 2008 through November 2012. January 2008 was chosen because several of the independent variables were measured in 2008 and prior to this time search volume was relatively low for global warming and climate change. We downloaded the data in November 2012, so that was the natural end point. For a fuller explanation of how and why we calculate GSI this way, please see the online appendix.

Google is mindful of not disclosing too much information. This is one reason why media markets are the smallest geographic unit at which data are available. Another important impact of Google's concern is that if search volume is lower than a threshold, which is undisclosed, then the AQS reported is zero. This effectively truncates GSI at zero. Weather is a popular search term, and there are no zeros for the 205 media markets. However, global warming and climate change are substantially less popular, and have 18 and 60 zeros, respectively. All media markets with zero AQS are still included in the analysis, but our regression models are designed to account for this truncation.

While we have scaled GSI relative to Washington DC in order to facilitate across space comparisons, the magnitudes of GSI values are difficult to interpret. To this end, we normalize the distribution of weather, global warming, and climate change GSI data to each have a standard deviation of 10. With this change, we can more easily interpret differences in GSI levels in terms of standard deviation. Further, when we present the regression analysis, we can interpret the coefficients in terms of standard deviation.

Table 1 presents summary statistics for the GSI of our three search terms: weather, global warming and climate change. We present means and standard deviations, as well as minimums and maximums. The distribution of GSI for global warming and climate change seems skewed, as in both cases the maximum is about six standard deviations greater than the mean. The means of the three GSI terms cannot be meaningfully compared. While the ordinal relationship is correct, with weather having the highest GSI followed by global warming and then climate change, the magnitudes of differences are incorrect. In reality, on a national scale, average query share for weather is 67 times higher than global warming, which is twice as large as climate change.²

3.2 Spatial trends in GSI

Using GIS, we created a map illustrating the spatial variation in GSI for each of our three search terms. These maps are displayed in Figure 1. The data are displayed classified by quintiles with darker areas indicating areas of high search interest. High levels of interest in weather are generally found in the higher latitudes of the country, while pockets in the Southern, Southwestern, and Eastern seaboard portions of the country show relatively little interest in weather. Trends for searches of global warming show high areas of interest centering in the interior portion of the country, particularly in the Midwest, Northwest and Great Lakes regions. Areas of low global warming interest exist particularly in the Southern and Northern Rocky Mountain regions. Similarly, searches for climate change show little interest in the Southern and Northern Rocky Mountain regions, while peak areas of interest are focused in the Southwest, West Coast and Northeast regions. Speculatively, high search interest in the Northeast could be a result of changing snowfall levels in that part of the country, and in the Midwest could be a direct result of major droughts in the corn and grain belt of the country.

Moving beyond simple spatial comparisons of GSI, we would like to identify statistically significant regions or clusters of interest and disinterest to inform policymakers where mitigation support may be high or low or prudent adaptation likely or unlikely. To this end, we completed a hotspot analysis of each Google variable using GIS, which identifies regions where several neighboring media markets have GSI that is a similarly large standard deviation from the

² These comparisons were made using the whole United States for the time period January 2008 through November 2012.

population mean. In order to determine those hot and cold regions, a Getis-Ord G_i^* statistic is calculated for each media market.³ For a hot (cold) spot to be designated, a media market must have a large positive (negative) G_i^* and be next to other media markets with similarly large G_i^* .

Figure 2 is the result of the hot spot analysis. The solid black and gray areas indicate hot spots, statistically significant clusters of high GSI, and the similarly colored hashed areas are cold spots, statistically significant areas of small GSI. The hot spot analysis for weather showed clusters of high values in the northern latitudes of the country, mostly in the regions east of the Rockies and west of the Great Lakes. Clusters of low interest in weather searches are found in the Southern portions of California and Nevada, much of New Mexico, Southern Texas and Florida, as well some parts of Appalachia. For global warming GSI, statistically significant clusters of high values were found in the Midwest, particularly south and west of the Great Lakes. Clusters of low values for global warming exist in the Southern region of the country stretching west from Georgia into central Texas, as well as an area encompassing much of Montana. Analysis for climate change searches show four distinct regional hotspots. A similar cluster of disinterest exists for climate change in the South, starting in central Texas and becoming more significant heading east into Louisiana and Mississippi, before losing significance over Georgia. Three major areas of climate change interest center around California,

³ The Getis-Ord local statistic is given as:

$$G_i^* = \frac{\sum_{j=1}^n w_{i,j} x_j - \bar{X} \sum_{j=1}^n w_{i,j}}{S \sqrt{\frac{n \sum_{j=1}^n w_{i,j}^2 - (\sum_{j=1}^n w_{i,j})^2}{n-1}}} \quad (1)$$

Where x_j is the attribute value for feature j , $w_{i,j}$ is the spatial weight between feature i and j , n is equal to the total number of features and:

$$\bar{X} = \frac{\sum_{j=1}^n x_j}{n} \quad (2)$$

$$S = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - (\bar{X})^2} \quad (3)$$

The G_i^* statistic is a z-score so no further calculations are required. The G_i^* statistic returned for each media market in the dataset is a z-score. For larger and statistically significant positive z-scores, intense the clustering of high values or a hot spot is indicated. For smaller and statistically significant negative z-scores, intense clustering of low values or a cold spot is indicated.

the Northeast, and Washington State. In two of these areas, regional CO₂ mitigation policies have been adopted (AB32 and RGGI).⁴

While it is unclear why these regions are hot and cold spots and what the determinants of high and low interest are, we feel these maps are of inherent value. Similar to the work that has been done in Development Economics with geographic targeting of aid for increased efficacy [30-31], these maps identify regions of high interest that can be harnessed for political support and areas of low interest that campaigns could target to rally support or may need targeted adaptation assistance. Areas of high interest have the potential for regional collaboration such as the efforts in California and New England described above. Conversely the areas of low interest indicate areas where awareness campaigns and increased community outreach may be required to inspire change. As an unfortunate example, the epicenter of Hurricane Katrina is an area of low interest. EPA and FEMA officials may want to target this area to raise awareness, increase understanding of impacts and encourage action in this region which is likely to be hard hit by the effects of climate change.

3.3 Covariates

To better explain the variation of GSI across media markets, we collected a series of variables that are intuitively pertinent to the motivation to seek information about climate change. Table 1 lists all the variables, their units, means and standard deviations, minimums and maximums. The sources of the data are as follows. Total population, percent white, percent over 65 years old and percent male come from the 2010 Census. Average income comes from the Bureau of Economic Analysis. CO₂ emissions per capita come from the Vulcan fossil fuel data project [32]. Democrat vote share is the share of votes received by the Democratic party in the 2008 Presidential Election, and these data come from Dave Liep's Atlas of U.S. Presidential Elections. Percent college graduate is the percent of the population that has completed college and was drawn from the five year American Community Survey 2006-2010. All of the above variables were downloaded at the county level. They were then aggregated to the media market level using a crosswalk between counties and media markets provided by Gentzkow and Shapiro

⁴ An analysis of temporal trends suggested that search interest did not change upon passage of either RGGI or AB32, which indicates that the results of the hotspot analysis are not a product of those policies.

[33]. Counties were weighted by population share within media market in order to attain the correct media market-wide measures.

We downloaded weather data from the National Oceanic and Atmospheric Administration's Global Historical Climatology Network for all years 2000-2011. For each day in the sample, we averaged all weather station data at the media market level. To estimate approximate climate, average annual temperature was calculated as the average temperature across all days and all years and annual total precipitation and annual total snow were calculated as the cumulative rain and snow for a given year, averaged over all years. Finally, coast is an indicator variable created using GIS, which equals one if a media market contains an ocean coastline.

4. Multivariate Analysis

In this section, we present the results of our multivariate regression analysis investigating the determinants of GSI for weather, global warming, and climate change. Table 2 presents the results. The first column shows the results for weather GSI and uses OLS. Because we have standardized each GSI term such that the standard deviation is 10.0, the coefficients are best interpreted in terms of the standard deviation.

The coefficient for weather searches on the coast variable is 2.80, which indicates that a coastal media markets on average have a GSI 0.28 standard deviations higher than inland media markets. The coefficients on average annual temperature and its square are -1.42 and 0.021, respectively, which indicate a convex quadratic function of temperature with a minimum at 33.8°C. This suggests that the coldest media markets have the highest search volume all else equal, and the volume declines at a decreasing rate as the climate warms for essentially the whole range of temperatures (only two media markets have an average temperature above 33.8°C). The significance of these climate variables on weather queries lends credence to the Google data and our modeling approach. However, the coefficients on precipitation and snow are insignificant. The results in Column 1 additionally show that several socioeconomic variables are significant determinates of weather GSI. Originally this idea is surprising, but when considering lifestyle patterns of different groups this result is understandable. For example, the coefficient on percent over 65 is 0.33 and statistically significant, indicating that metropolitan areas with larger senior populations are more likely to seek information about the weather. With the caveat of ecological inference, this may suggest that seniors, who may be more affected by adverse

weather than others, are more likely to search for weather on the internet. The coefficient on Democrat vote share is negative and significant, suggesting the Democratic leaning metropolitan areas are less likely to search weather. The coefficient on log population is negative and significant, indicating that metropolitan areas with smaller populations are more likely to search for weather. This potentially could be a result of a larger agricultural sector in low population areas, less infrastructure to deal with adverse weather, or simply less access to alternative forecasts from, say, local news. These results point out that people with varying socioeconomic characteristics may also care about weather in different ways. This evidence solidified our intuition that socioeconomic variables may be important determinants of global warming and climate change GSI.

Columns 2 and 3 display the results of Tobit regressions on global warming and climate change GSI. We use a Tobit regression because a number of our dependent variables were censored by Google Trends because – for privacy reasons – GSI will record as zero if a media market does not have enough search volume to reach a minimal threshold.⁵ For global warming GSI, income, percent college graduate, percent white, and percent over 65 are statistically significant. For every thousand dollar increase in media market income, global warming GSI decreases by 0.08 standard deviations. As the percent of college graduates increases one percent, global warming GSI increases 0.061 standard deviations. Global warming GSI increases by 0.021 and 0.065 standard deviations as the share of whites and share of over 65 increases by one percent, respectively. The coefficient on log population is also significant, but this is likely a function of Google’s volume cutoff for disclosure and small media market’s inability to have search volume above that mark. Surprisingly, geographic and climate variables are not correlated with search interest.

Column 3 examines climate change GSI and finds that the determinants are similar to those of global warming GSI. Income has a negative effect and share white has a positive effect on climate change GSI. As with global warming GSI, percent college graduate has a significant, positive effect, but for climate change GSI the magnitude is 2.6 times larger, which suggests that climate change is a more common search term in well-educated metropolitan areas. A more substantial difference is that Democrat vote share is a significant determinant of climate change

⁵ While Google doesn’t truncate at zero, the index that is reported by Google Trends is truncated at zero and that is why we use zero as the truncation point. The interpretation of each coefficient is in terms of marginal changes in the latent (uncensored) variable.

GSI; a one percent increase in Democrat vote share raises GSI by 0.028 standard deviations. This is intuitive as Democrats tend to be more supportive of climate change science and policies.⁶

4.1 Relationship between climate change GSI and global warming GSI

Many people think of global warming and climate change as synonymous, and certainly the results of Table 2 suggest a similar set of determinants. However, we wanted to examine if there are differences in the types of metropolitan areas that would search one term over the other.

Table 3 gives the results of a model with climate change GSI as the dependent variable and global warming GSI as an explanatory variable, along with the full set of covariates from Table 2.

The coefficient on global warming GSI indicates that while these two measures of GSI are strongly positively correlated, the correlation is far from perfect. Democrat vote share, percent college graduate and percent white are still significant determinants of climate change GSI even when controlling for global warming GSI. These results suggest that there could be socioeconomic divides in the types of people that choose to search for climate change versus global warming. Specifically, metropolitan areas that are more liberal, more educated and more white tend to search for climate change. The education result seems intuitive as climate change is a more technical term, and people living in metropolitan areas with a highly educated population would be more likely to be exposed to this term.

4.2 Is there an optimal mix of political ideology for engagement?

While the results so far have shown a positive correlation between Democrat vote share and GSI, given the prominence of politics in debates over climate change, we chose to examine this relationship in more detail. We first obtained residuals from a regression of GSI on all covariates except Democrat vote share, to partial out the correlation due to other covariates. Then we created non-parametric plots of those residuals against Democrat vote share. These results are given in Figure 3. The top panel, which displays the plot for global warming GSI, shows that

⁶ As an additional validity check, we downloaded data for the search term “Mars”, which we view as a scientific search term, but without much if any political bend. If the determinants of Mars GSI were similar as those for global warming GSI and climate change GSI, it would question whether the results were specific to climate change or just search terms in general. The results indicate that only income and log population are determinants of Mars GSI, suggesting that global warming/climate change is somewhat unique and using Google Trends may lead to important insights about spatial disparities in engagement.

GSI increases steeply as metropolitan areas become less Republican, up until around a Democrat vote share of 43%, at which point GSI begins declining gradually. This inverse-U shaped relationship explains why Democrat vote share was an insignificant determinant of global warming GSI in Table 2. For climate change GSI, displayed in the bottom panel, the relationship is always increasing, indicating that climate change GSI increases as metropolitan areas lean more Democratic over the whole range of the political spectrum. These results first suggest that climate change is more of a branded political word than global warming. Second, with global warming, our results indicate that metropolitan areas that are more moderate are seeking information and engaging in this topic, perhaps implying healthy debate.

5. Conclusions

This paper uses a novel data set of geographically explicit online information searches about climate change to assess the extent of geographic clustering of search interest and the climatological and socioeconomic determinants of search interest. Additionally, the data are distinctly private and anonymous, which may curtail any posturing that occurs during surveys about this divisive issue. Our results identify regions and the characteristics of metropolitan areas that actively seek information. One goal with these results is to inform the political process of passing meaningful carbon mitigation legislation by determining which areas of the United States and which types of cities are likely to support legislation and which will need to be campaigned for. Regardless of future legislation, some impacts of climate change are inevitable. Our results additionally identify the types of metropolitan areas that may adapt better to changes due to their knowledge and also which metropolitan areas and regions may need assistance.

In light of the findings of this paper, there are several areas of future research that would be valuable. Ultimately, one of the broader goals of this and other similar papers is to understand what shapes public opinion in order to pursue meaningful mitigation policies. This paper measures information seeking behavior, but this is fundamentally different than an opinion. Future research should seek to assess how the process of acquiring knowledge matters for policy support. Also, this paper measures static differences in information seeking behavior. It is just as critical to understand how behavior changes over time and what types of areas are most likely to change their behavior. On the adaptation side, future research could examine how information

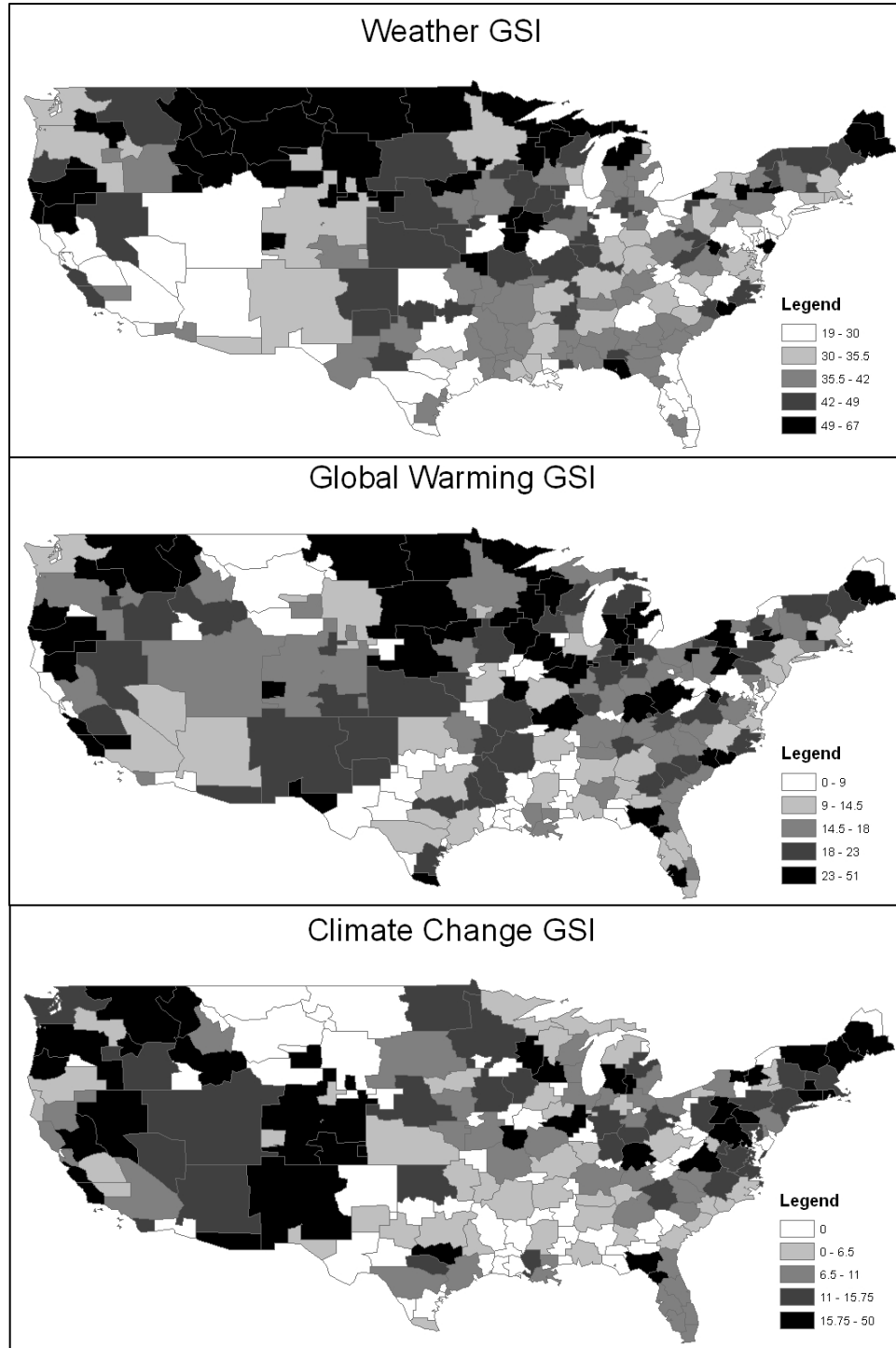
acquisition affects decisions about energy efficiency choices or flood protection or a host of smart adaptation strategies.

References

1. Kahn, M.E. (2010). *Climatopolis: How Our Cities Will Thrive in the Hotter Future*. Basic Books, New York, NY.
2. O'Connor, R.E., Bord, R.J., Yarnal, B., & Wiefek, N. (2002). Who wants to reduce greenhouse gas emissions? *Social Science Quarterly*, 83(1), 1–17.
3. Akerlof, K., Maibach, E.W., Fitzgerald, D., Cedeno, A.Y., & Neuman, A. (2013). Do people 'personally experience' global warming, and if so how, and does it matter? *Global Environmental Change*, 23, 81-91.
4. Egan, P.J., and Mullin M. (2012). Turning personal experience into political attitudes: the effect of local weather on Americans' perceptions about global warming. *The Journal of Politics*, 73(3), 796-809.
5. Dunlap, R.E., & McCright, A.M. (2008). A widening gap: view on climate change. *Environment*, 50(5), 26-35.
6. Krosnick, J.A., Holbrook, A.L., Lowe, L., & Visser, P.S. (2006). The origins and consequences of democratic citizens' policy agendas: a study of popular concern about global warming. *Climatic Change*, 77, 7-43.
7. Shwom, R., Bidwell, D., Dan, A., & Dietz, T. (2010). Understanding U.S. public support for domestic climate change policies. *Global Environmental Change*, 20, 472-82.
8. Borick, C.P., & Rabe, B.G. (2010). A reason to believe: examining the factors that determine individual view on global warming. *Social Science Quarterly*, 91(3), 777-800.
9. Brulle, R.J., Carmichael, J., & Jenkins, J.C. (2012). Shifting public opinion on climate change: an empirical assessment of factors influencing concern over climate change in the U.S., 2002-2010. *Climatic Change*, 114, 169-88.
10. Hamilton, L.C., & Keim, B.D. (2009). Regional variation in perceptions about climate change. *International Journal of Climatology*, 29, 2348-52.
11. Purcell, Kristen, Brenner, Joanna, and Rainie, Lee (2012) *Search Engine Use 2012*, Washington, DC: Pew Internet and American Life Project.
12. Scheitle, Christopher P. (2011). Google's Insights for Search: A Note Evaluating the Use of Search Engine Data in Social Research. *Social Science Quarterly*, 92(1), 285-95.
13. Reilly, Shauna, Sean Richey, and J. Benjamin Taylor. "Using Google Search Data for State Politics Research An Empirical Validity Test Using Roll-Off Data." *State Politics & Policy Quarterly* 12, no. 2 (June 1, 2012): 146–59.
14. Sherman-Morris, Kathleen, Senkbeil, Jason, and Carver, Robert (2011) Who's Googling What? What Internet Searches Reveal about Hurricane Information Seeking. *Bulletin of the American Meteorological Society*, 975-85.
15. Choi, H., & Varian, H. (2012). Predicting the present with Google Trends. *The Economic Record*, 88, 2-9.
16. Hand, Chris, and Judge, Guy (2012). Searching for the picture: forecasting UK cinema admissions using Google Trends data. *Applied Economics Letters* 19, 1051-55.

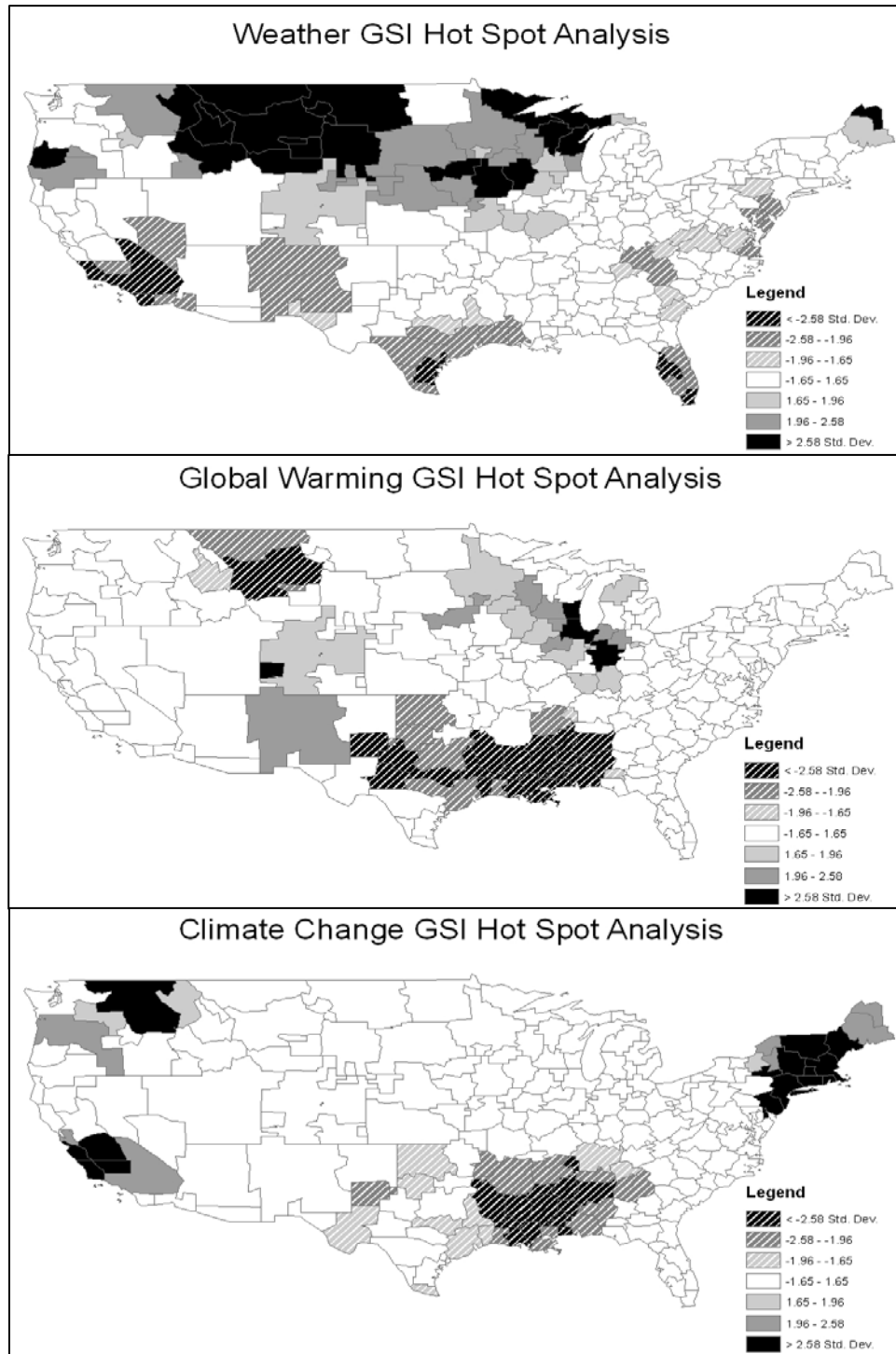
17. Ginsberg, J., Mohebbi, M.H., Patel, R.S., Brammer, L., Smolinski, M.S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, 457, 1012-14.
18. Pelc, Krzysztof J. (2013). Googling the WTO: What Search Engine Data Tell Us About the Political Economy of Institutions. *International Organization*, 67(3), 629-55.
19. Stephens-Davidowitz, S. (2012). The effects of racial animus on a black presidential candidate: using Google search data to find what surveys miss. Harvard University Working Paper.
20. Kahn, M.E., & Kotchen, M.J. (2011). Business cycle effects on concern about climate change: the chilling effect of recession. *Climate Change Economics*, 2(3), 257-73.
21. Lang, C. (2014) "Do weather fluctuations cause people to seek information about climate change?" *Climatic Change*, 125, 291-303.
22. Mellon, Jonathan (2013). Where and When Can We Use Google Trends to Measure Issue Salience?" *PS: Political Science & Politics* 46(2), 280-90.
23. Swearingen, C. Douglas, and Ripberger, Joseph T. (2014). Google Insights and U.S. Senate Elections: Does Search Traffic Provide a Valid Measure of Public Attention to Political Candidates? *Social Science Quarterly*, n. pag.
24. Manzano, Sylvia, and Ura, Joseph (2013). Desperately Seeking Sonia? Latino Heterogeneity and Geographic Variation in Web Searches for Judge Sonia Sotomayor. *Political Communication*, 30(1), 81-99.
25. Scharhow, Michael, and Vogelgesang, Jens (2011). Measure the Public Agenda using Search Engine Queries. *International Journal of Public Opinion Research*, 23(1).
26. Carrière-Swallow, Yan, and Labbé, Felipe (2013). Nowcasting with Google Trends in an Emerging Market. *Journal of Forecasting*, 32, 289-98.
27. Carneiro, Herman, and Mylonakis, Eleftherios (2009) Google Trends: A Web-based Tool for Real-Time Surveillance of Disease Outbreaks. *Clinical Infectious Disease*, 49(10).
28. Rose, Daniel E., and Levinson, Danny (2004). Understanding User Goals in Web Search. *Proceedings of the 13th International Conference on World Wide Web*. New York: ACM, n.d. 13-19.
29. Ripberger, Joseph T. (2011). Capturing Curiosity: Using Internet Search Trends to Measure Public Attentivenessps. *The Policy Studies Journal*, 39(2), 239-59.
30. Baker, J., & Grosh, M. (1994). Poverty reduction through geographic targeting: How well does it work? *World Development*, 22(7), 983-995.
31. Lang, C., Barrett, C., & Naschold, F. (2013). Targeting maps: an asset-based approach to geographic targeting. *World Development*, 41(1), 232-244.
32. Gurney, K.R., Mendoza, D., Zhou, Y., Fischer, M., Miller, C., Geethakumar, S., de la Rue du, S. (2009) The Vulcan Project: High resolution fossil fuel combustion CO2 emissions fluxes for the United States, *Environmental Science and Technology*, 43(14), 5535-5541.
33. Gentzkow, M. and Shapiro, J.M. (2008). Introduction of Television to the United States Media Market, 1946-1960. ICPSR22720-v1. Chicago, IL: University of Chicago/Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributors], 2008-09-30. doi:10.3886/ICPSR22720.v1

Figure 1: Spatial Distribution of Google Search Index



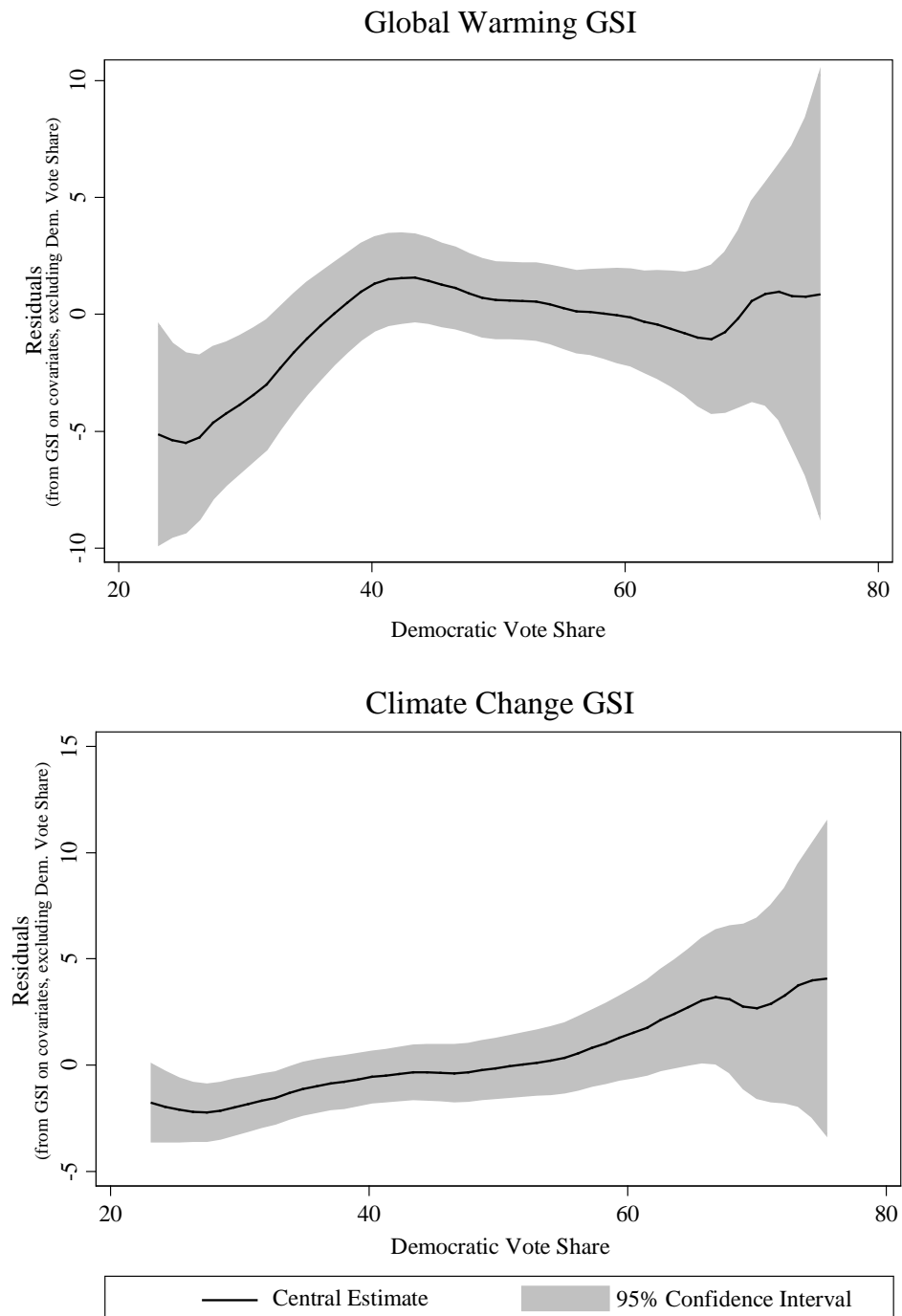
Notes: Darker designated market area's (DMA's) show regions where there is a higher interest level (compared to Washington D.C.) and the lighter DMA's indicate less interest. The colors correspond to the average percentage of interest level compared to the maximum over a four year period. The data is represented in quantiles with twenty percent of data in each interval.

Figure 2: Hot Spot Analysis



Notes: Hot spots, statistically significant clusters of high GSI values are in solid black and grey. Cold spots, statistically significant clusters of low GSI values are in hash marked grey and black. Clusters are evaluated based upon calculated G_i^* statistics and associated z-scores and p-values (See footnote 3 for a full explanation).

Figure 3: The Nonparametric Relationship of Democratic Vote Share and GSI



Notes: Residuals are obtained from a regression of GSI on all control variables shown in Table 2 except for Democrat vote share. The residuals are then plotted against Democratic vote share. Stata suggested bandwidths and local mean smoothing are used.

Table 1: Summary Statistics

Variable	Mean(Std Dev)	Min	Max
GSI Weather	40.66(10.0)	19.15	67.03
GSI Climate Change	8.61(10.0)	0	49.92
GSI Global Warming	16.21(10.0)	0	51.44
Income (1000\$)	35.89(5.76)	20.47	59.87
Population (1000 persons)	1479(2370)	7.69	20852
CO ₂ Emissions (tons per capita)	7.12(7.28)	1.93	72.30
Democrat Vote Share	47.45(10.47)	23.17	75.34
Percent College Graduate	17.75(4.96)	8.23	32.37
Percent White	79.40(12.43)	34.36	97.57
Percent Over 65	14.10(2.49)	7.93	25.11
Percent Male	49.40(0.79)	47.64	52.17
Coast (1=yes)	0.20(0.40)	0	1
Average Annual Temperature (Celsius)	16.16(5.86)	4.66	41.30
Annual Total Precipitation (inches)	52.31(19.17)	4.12	99.79
Annual Total Snow (inches)	20.75(23.57)	0	113.42

Notes: GSI is gathered from Google Trends and is derived from raw Google search data generated by individuals across the United States from January 2008 through November 2012. We normalized the standard deviation to 10 to allow the spread of the distribution of a variable with a large mean and correspondingly large standard deviation to be compared more appropriately with the spread of the distribution of another variable with smaller mean and correspondingly smaller standard deviation. Average income data comes from The Bureau of Economic Analysis. Carbon dioxide emissions were gathered from the Vulcan fossil fuel CO₂ emissions data, produced at Arizona State University in conjunction with collaborators at Colorado State University and Lawrence Berkeley National Laboratory. The Democrat data is from Dave Liep's Atlas of U.S. Presidential Elections 2008. College graduate data is from the 5 year American Community Survey from 2006 until 2010. The climate variables were gathered daily from the National Oceanic and Atmospheric Administration's Global Historical Climatology Network.

Table 2: Determinants of Weather, Global Warming and Climate Change GSI

Independent variables	(1) Weather	(2) Global Warming	(3) Climate Change
Income (1000\$)	-0.11 (0.09)	-0.80*** (0.17)	-0.67*** (0.18)
Democrat Vote Share	-0.13*** (0.04)	0.10 (0.08)	0.28*** (0.09)
Percent College Graduate	0.17 (0.11)	0.61*** (0.19)	1.62*** (0.21)
Percent White	-0.0079 (0.04)	0.21** (0.08)	0.34*** (0.09)
Percent Over 65	0.33* (0.19)	0.65* (0.349)	-0.10 (0.37)
Percent Male	0.88 (0.61)	0.22 (1.10)	-0.50 (1.19)
Population (1000 persons) (log)	-4.83*** (-0.43)	4.63*** (0.80)	4.48*** (0.88)
CO ₂ Emissions (tons per capita)	0.01 (0.05)	-0.01 (0.10)	-0.10 (0.14)
Coast (1=yes)	2.80** (1.18)	1.56 (2.09)	-1.61 (2.16)
Average Annual Temperature (Celsius)	-1.42*** (0.36)	-0.56 (0.65)	-0.30 (0.68)
Average Annual Temperature (Celsius) squared	0.021*** (0.0082)	0.0040 (0.015)	-0.00027 (0.016)
Annual Total Precipitation (inches)	-0.014 (0.025)	-0.061 (0.045)	0.076 (0.048)
Annual Total Snow (inches)	0.021 (0.027)	-0.041 (0.048)	-0.028 (0.049)
Number of left-censored observations at GSI≤0	0	18	60
$\rho_{y,\hat{y}}^2$	0.75	0.52	0.66

Notes: Column (1) uses OLS. Columns (2) and (3) use a Tobit regression. The sample size is 205 for all columns. Standard errors are in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01

Table 3: Model of GSI of Climate Change search on Global Warming search

Independent variables	
GSI Global Warming	0.41*** (0.11)
Democrat Vote Share	0.23*** (0.085)
Income (1000\$)	-0.29 (0.18)
Percent College Graduate	1.30*** (0.21)
Percent White	0.20** (0.097)
Person Over 65	-0.37 (0.37)
Percent Male	-1.49 (1.25)
Population (1000 persons) (log)	1.24 (0.95)
CO ₂ Emissions (tons per capita)	-0.019 (0.17)
Coast (1=yes)	-2.00 (2.03)
Average Annual Temperature (Celsius)	-0.37 (0.65)
Average Annual Temperature (Celsius) squared	0.0029 (0.015)
Annual Total Precipitation (inches)	0.075 (0.049)
Annual Total Snow (inches)	-0.01 (0.045)
Sample Size	145
$\rho^2_{y,\hat{y}}$	0.47

Notes: The results shown come from an ordinary least squares *regression* with climate change GSI as the dependent variable. The Data with 0 in GSI of Climate Change or Global Warming are removed, which results in number of observations being 145. Standard errors are in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01